

AD-A265 144



(11)

Interim Tile Design Study  
for the  
Digital Chart of the World

Contract No.: DMA600-89-C-0023  
CDRL B001

Date Prepared: July 1990

Prepared for:  
Defense Mapping Agency Systems Center  
12100 Sunset Hills Road, Ste. 200  
Reston, Virginia 22090-3207

DTIC  
ELECTE  
MAY 27 1993  
S E D

Attn: Ms. Lynn Martin  
Contract Specialist

Prepared by:  
Environmental Systems Research Institute, Inc.  
380 New York Street  
Redlands, California 92373  
U.S.A.

STRIKED STATEMENT  
Approved for public release  
Distribution Unlimited

93 5 20 00 2

426588

93-11831



## **Introduction**

The Digital Chart of the World (DCW), a global digital database suitable for geographic analysis, is estimated to require 15 gigabytes of storage. Given that few storage devices or processors currently available can maintain a database this size as a single unit, it will be necessary to geographically partition the database. To permit use of the storage medium selected for the DCW (550-megabyte CD-ROM diskettes), the DCW will have to be divided into 25 or more partitions, each containing data for a portion of the world. In addition to the geographic partitioning necessitated by storage on CD-ROMs, however, further geographic partitioning (into small units or tiles) is being studied, because tiling, if appropriately done, may improve both the efficiency with which the database can be accessed and the ease with which it is used.

This interim report documents ESRI's continuing investigation into tiling for the DCW. The first part of the paper reviews the characteristics of tiling schemes in general. The second identifies requirements and goals for DCW tiling in particular and summarizes project experience with tiling to date.

## **Tiling in the Digital Environment**

### **Impact of Data Storage Characteristics**

A need to break a large geographic area into smaller, more manageable units accompanies almost any mapping effort that covers a large area of the globe. A conventional paper map series is generally made up of maps showing data for a given theme or collection of complementary themes. The constraints on physical map-making include those connected with compilation, printing, storage, and distribution. The size and scale of individual sheets are controlled by the level of detail collected for the theme or themes in conjunction with the aesthetic requirements of presentation. Since the global distribution of thematic features varies, different amounts of data of a single type are present on every sheet. For example, a map of central Canada contains an abundance of hydrographic features, while a map of an area of similar size in northern Africa does not.

In a digital environment, storage is defined in terms of file size in bytes, rather than length and width. The amount of data that is most appropriate for an area from the storage standpoint is not constrained by the amount of information that can be clearly displayed on an individual map sheet. However, the most appropriate amount of digital spatial data is constrained by such things as the size and availability of storage units, the organization of the database and the method of database access. To address these complex system constraints on the management of geographic data, large digital databases are typically subdivided into an organized collection of units of data called tiles.

### **Impact of Software Characteristics**

What layering is to thematic partitioning, tiling is to spatial partitioning. More formally, tiling is the logical and physical division of a spatial database into a collection of geographically manageable units. From a computer system perspective, the most appropriate size for a single unit can be determined by evaluating the interactions between secondary storage devices, a set of known computer parameters (e.g., minimum hardware configuration), defined software functions, and user requirements. The most appropriate scheme for an entire set of tiles or tiling scheme can be determined by comparing the performance of these schemes. We are evaluating two basic types of tiling schemes. In the first, the area of the tile is fixed and the amount of data in each tile is allowed to vary. In

2 July 1990

Interim Study, Jan 90 - July 90

Interim Title Design Study for the Digital Chart  
of the World

C: DMA600-89-C-0023

Environmental Systems Research Institute, Inc.

Environmental Systems Research Institute, Inc.  
380 New York Street  
Redlands, CA 92373

Defense Mapping Agency  
8613 Lee Highway  
Fairfax, VA 22031-2137

DMA/TR--93/02

Distribution Unlimited

This study reviews investigations for the partitioning of large spatial databases undertaken in the development of the Defense Mapping Agency's new vector product the Digital Chart of the World (DCW). The DCW is a geo-spatial, topologically structured, global database designed for use in Geographic Information Systems (GIS). The study covers the characteristics of tiling, or partitioning, of large digital geo-spatial databases to improve data storage, access, and use. It identifies tiling requirements and goals for the DCW and summarizes project experience. It is the first of two reports performed under the DCW development effort. The second is the Final Tile Design Study for the Digital Chart of the World, November 1990.

Geo-spatial database partitioning, tiling, Geographic Information System, GIS, spatial databases, Vector Product Format, VPF, Digital Geographic Information Exchange Standard, DIGEST, Geographic Information Standards, digital cartography, digitizing, Geo-spatial Information, CD-ROM performance, Defense Mapping Agency, DMA, Vector Maps, Digital Chart of the World, DCW.

20

Unclassified

Unclassified

Unclassified

UL

the second, the amount of data in each tile is approximately equal and the area of the tile is allowed to vary. The objective of the assessment is to achieve predictability and, with that, to acquire control. The greater the number of controlled components in the database, the greater the assurance of proper functionality.

## Tile Models Reviewed

Five planar and one nonplanar tile model have been reviewed for possible use with the DCW; nonsystematic fixed, quasi-systematic fixed, systematic fixed, irregular adaptive, regular adaptive, and recursion of a solid. These models can be categorized according to their characteristics. A tiling scheme that, once defined, does not allow any alteration to the number of spatial divisions is termed fixed. Fixed tiling schemes are often utilized for stable, well-defined applications or for the distribution of digital data to nonspecific audiences. Fixed tiling schemes can be further categorized as nonsystematic, quasi-systematic, or systematic. Adaptive tiling schemes, in contrast to fixed tiling schemes, are those that allow changes to the spatial divisions. Adaptive tiling schemes may be further defined as either regular or irregular. Figure 1 illustrates this tiling classification hierarchy, and Figure 2 presents graphic examples of each.

### Fixed Tiling

**Nonsystematic Tiling.** Fixed tiling schemes having spatial extents not defined along consistent lines of the graticule are termed nonsystematic. In the example shown in Figure 2(a), the tiling boundaries may follow physiographic features, such as drainage basins, or some administrative units. This tiling scheme is appropriate for applications that require data to be managed according to a predefined classification system, such as the U.S. Census Bureau's TIGER database (Broome and Meixler, 1990), which is tiled along county lines, and the Level 3-C DFAD database (*Product Specifications for Digital Feature Analysis Data*, 1988). The Level 3-C DFAD vector data is partitioned into tiles that vary in size from 2 by 2 nautical miles to 10 by 10 nautical miles; the tiles may preserve equal geographic extents of coverage for all parts of the globe but do not consistently follow lines of longitude and latitude. As a result, the spatial relationships of these tiles (the distances from each other, directions to each other, and neighbors of each) must be determined by looking at the coordinates making up their boundaries or by searching an index.

**Quasi-Systematic Tiling.** The boundaries in quasi-systematic tiling scheme follow lines of the graticule but at inconsistent intervals. The partitioning of the 1:1,000,000-scale ONC map series into a set of rectangular sheets (*Standard Index Chart*, 1988; Figure 2(b)) is an example of quasi-systematic fixed tiling. Although the ONC sheets are referenced to the graticule in whole-degree units, the longitudinal extent of the ONC sheets varies with latitude. In other words, although the ONC sheets always cover 8° bands of latitude, they cover increasing extents of longitude towards the poles. With this tiling scheme, an algorithm can be used to access tiles along lines of latitude, but a direct examination of the coordinates defining boundary edges is necessary to determine the positions of the map sheets with respect to longitude. All map sheets with the letter designation 'L' will be within the latitude interval from 0° to 8° north and will be 24° in latitude from map sheets designated with letter 'Q'. This capacity for internal referencing is referred to as providing an implicit spatial relationship, since it is inherent in the tile scheme. The longitudinal relationships among map sheets, however, even along the same interval of latitude, must be determined either through direct examination of the map sheet boundaries or through an index. In addition, since the ONC series covers only land masses and there are gaps in the

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

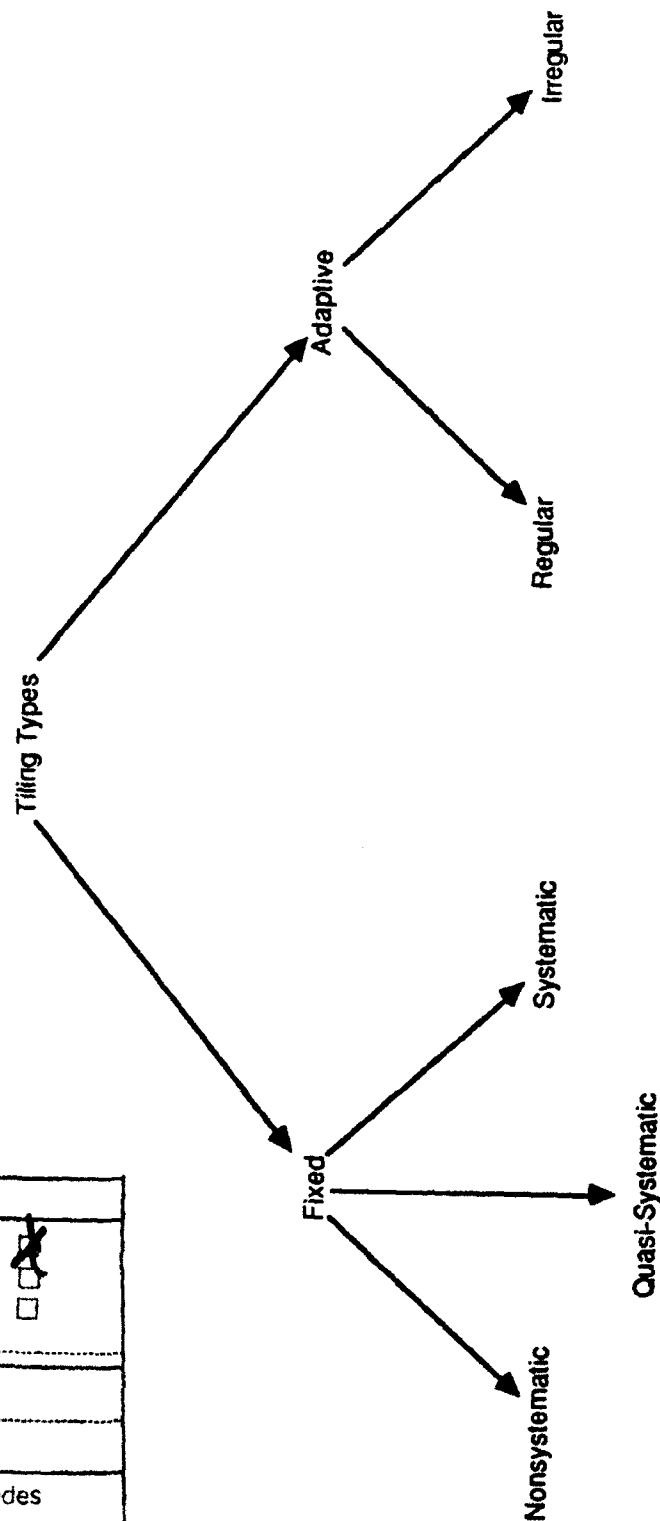

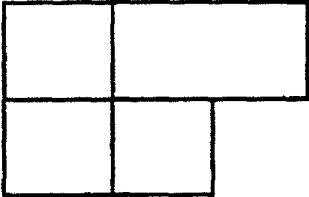
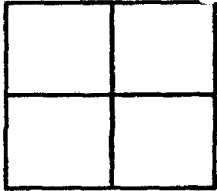
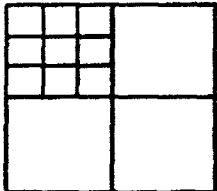
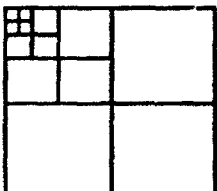


Figure 1. Relationship of Tiling Types.

Graphic	Type	Example
 <p>(a)</p>	Nonsystematic, fixed	TIGER DFAD, Level 3-C
 <p>(b)</p>	Quasi-systematic, fixed	ONC series sheets
 <p>(c)</p>	Systematic, fixed	DTED DFAD, Level 1-C
 <p>(d)</p>	Systematic, adaptive, irregular	GEOREF
 <p>(e)</p>	Systematic, adaptive, regular	Quadtree approach

**Figure 2. Tiling Schemes.**

indexing scheme coverage (Figure 3), the ONC series does not in and of itself provide global tiling coverage.

**Systematic Tiling.** A fixed tiling scheme in which the spatial extents of the tiles follow consistent lines of latitude and longitude is termed systematic. Level 1 DTED (*Product Specifications for Digital Terrain Elevation Model*, 1986) and Level 1-C DFAD (Figure 2(c)) are examples of systematic fixed tiling schemes. The point data in Level 1 DTED and the vector data in Level 1-C DFAD are distributed in consistent  $1^\circ$  by  $1^\circ$  cells. The Level 2 DTED that is distributed in  $1^\circ$  by  $1^\circ$  spatial extents is also systematic. The Level 2 DTED that is distributed in less than  $1^\circ$  by  $1^\circ$  cells is in either a quasi-systematic or a nonsystematic partitioning scheme.

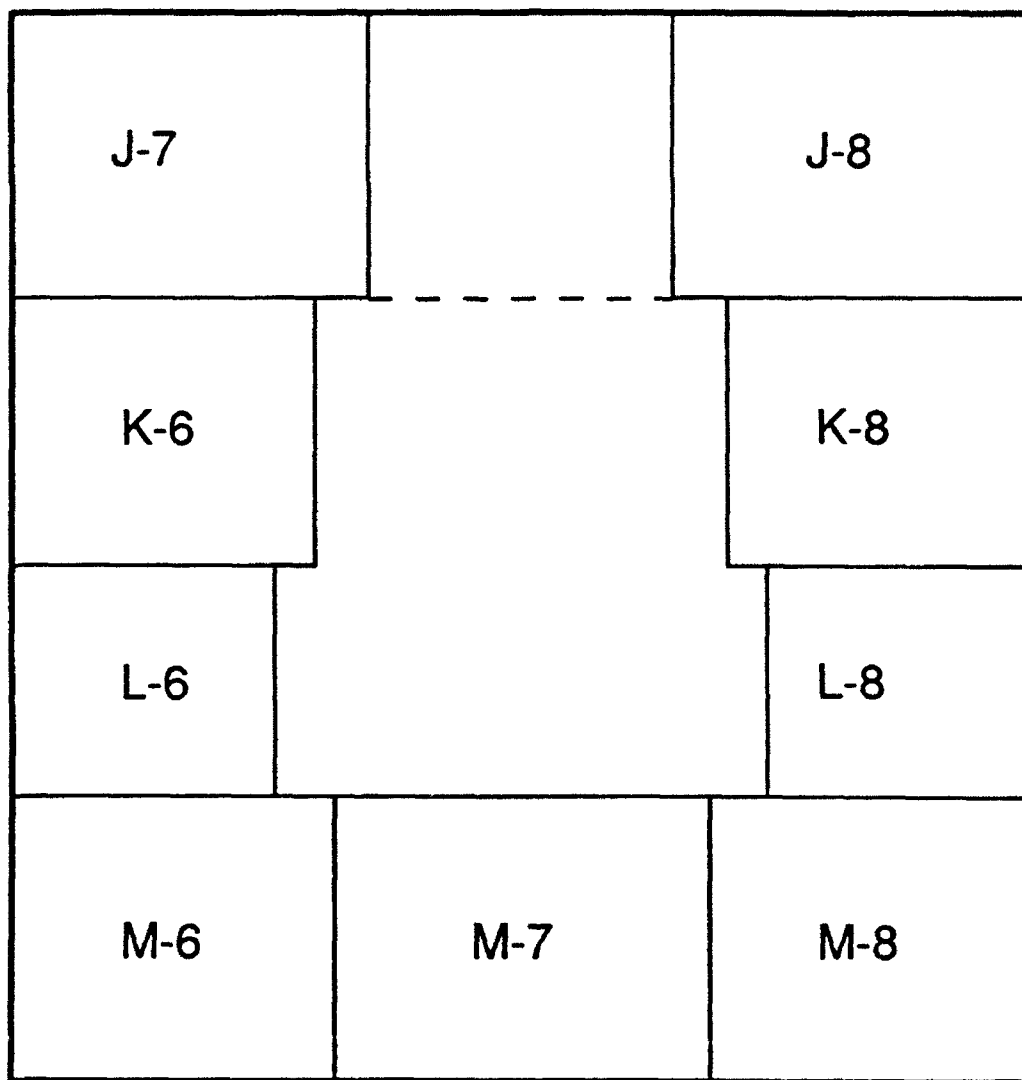
### Adaptive Tiling

The second general category of tiling schemes, termed adaptive, allows tile boundaries to change in response to boundary determination criteria. In contrast to the fixed tiling schemes discussed previously, adaptive tiling schemes permit geographic coverage to be modified according to a criterion such as a targeted file size. A comparison of the interaction between tile size and data density for fixed and adaptive tiling schemes is illustrated in Figure 4. The illustration on the left shows a set of fixed tiles; the amount of data varies per tile. The illustration on the right shows a set of tiles after they have been modified, or adapted, to meet a hypothetical file size requirement of approximately 25 megabytes. Most adaptive tiling schemes modify the spatial extent of the tiles through a process of hierarchical division (Figure 2(d)), which differs from trial and error in that it is performed systematically (the largest tile in the hierarchy is developed along consistent lines of longitude and latitude). Hierarchical division is presumed in the following discussion.

**Irregular Adaptive Tiling.** Adaptive hierarchical tiling can be either regular or irregular. Irregular tiling is a system of spatial partitioning in which the same number of divisions is not used at all levels. The World Geographic Reference System (GEOREF; see *Datums, Ellipsoids, Grids and Grid Reference Systems*, 1986), which is illustrated in Figure 5, is an example of graticule-based irregular adaptive tiling. Partitioning within GEOREF does not proceed in a consistent manner with depth. At the highest level of generalization, the earth is divided into 288  $15^\circ$  by  $15^\circ$  quadrangles. At the next level, each of the 288 initial divisions is divided into 225  $1^\circ$  by  $1^\circ$  quadrangles, which are subsequently divided into 3600  $1'$  by  $1'$  quadrangles.

Although the tiling for GEOREF is irregular, it illustrates three properties that would be valuable for the DCW tiling scheme: 1) it is global, covering  $360^\circ$  of longitude and  $180^\circ$  of latitude; 2) it is hierarchical, so selected areas can be further subdivided; and 3) it uses an addressing scheme compact enough to permit a combination of four letters and four numbers to reference any point on earth to within 1 nautical mile of its actual position. Although the addressing is compact, however, it does require the use of two numbering systems (base 26 for the letters used at the two highest levels of division and base 60 for the two subsequent degree and minute divisions).

**Regular Adaptive Tiling.** Regular tiling schemes use a system of division that is constant (recursive) at all levels. Unlike irregular division, recursive division allows for addressing on a single base numbering system, and in an adaptive tiling scheme the division can be designed to stop at different levels in different areas. Figure 2(e) illustrates



**Figure 3. Example of a Gap in ONC Coverage.**



Variable tile size, approximately equal data quantity

25.5	23	25
24	25	25
25	26	27

Fixed tile size, variable data quantity

3	30	18	8	10
0	10	15	25	25
0	5	50	15	7

Figure 4. Interaction Between Tile Size and Data Quantity. (Data Quantities Are in Megabytes.)

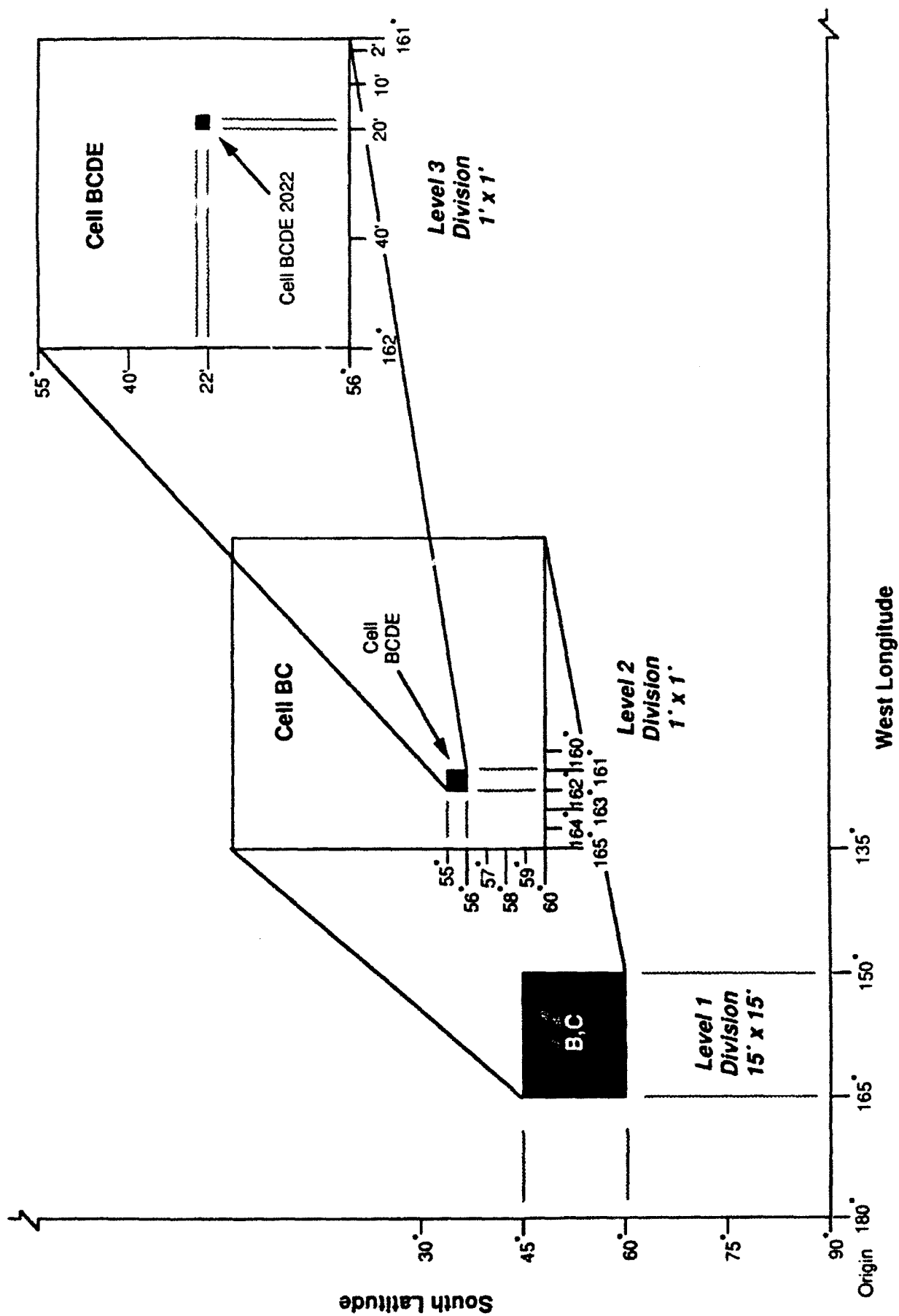


Figure 5. GEOREF Addressing: An Irregular Nested Grid System.

a system of recursion by even divisions of four to yield a quadtree model. In Figure 6, the top graph illustrates the tree structure that results from selective recursive division of an area. The lower graph in Figure 6 shows how that division would appear spatially. (In the tree structure 'Y' indicates a decision to continue subdividing and 'N' indicates that the value of data content is acceptable). The data quantity for the largest area, Cell A, exceeds a targeted maximum, and so Cell A is therefore evenly divided into fourths, labeled 0, 1, 2 and 3. The numbering scheme is consistent, beginning in the upper left-hand corner. In this example, the data in the areas addressed 0 and 3 still exceeds the targeted maximum, so these areas are again subdivided into fourths. After the second division, the upper left quarter of area 0, addressed 00, and the lower right quarter of area 3, addressed 33, still exceed the data quantity maximum. Each of these areas will again be quartered, at which point the data quantity for all areas is less than the targeted maximum and the recursive division stops. The address of the shaded area in the lower part of Figure 6, A330, indicates not only the area's relative position within Cell A, but also, by its length, the relative amount of data in any area; the more dense the data, the longer the address.

Another advantage of regular adaptive tiling is that the tree structure that characterizes it can be used as the structure for a location for storing local data dictionaries. In other words, an attribute table can be developed at each of the nodes that identifies the features, attributes and attribute values of the areas (branches) beneath it. These tables, although they increase overhead, can be useful when one is selectively copying from the CD.

### Nonplanar Tiling

All the tiling schemes discussed to this point assume a plane surface in decimal degrees. Each will therefore exhibit the distortion produced when spherical coordinates are represented on a plane (*Which Map Is Best?*, 1986). Of particular interest to tiling schemes is the difference in area that accompanies changes in latitude. Table 1 (from Robinson and Sale, 1969) lists the changes in area of  $1^\circ$  by  $1^\circ$  cells with changes in latitude.

**Table 1. Approximate Areas of  $1^\circ \times 1^\circ$  Cells by Latitude**

Latitude (deg)	Area (sq. mi.)	Area (sq. km)
0-1	4,752	12,308
10-11	4,674	12,106
20-21	4,459	11,549
30-31	4,109	10,642
40-41	3,634	9,412
50-51	3,047	7,892
60-61	2,364	6,123
70-71	1,605	4,157
80-81	795	2,059
88-89	42	109

Source: Robinson and Sale (1969).

Some new models that rely on a solid, or polyhedron, to represent the earth have been developed to reduce the amount of distortion introduced by planar schemes. Five candidate solids have been cited in the literature (Peuquet, 1988); they correspond to the five platonic solids (the dodecahedron, tetrahedron, cube, octahedron, and icosahedron). If the desirable properties of an adaptive, systematic, regular tiling scheme that apply to planar

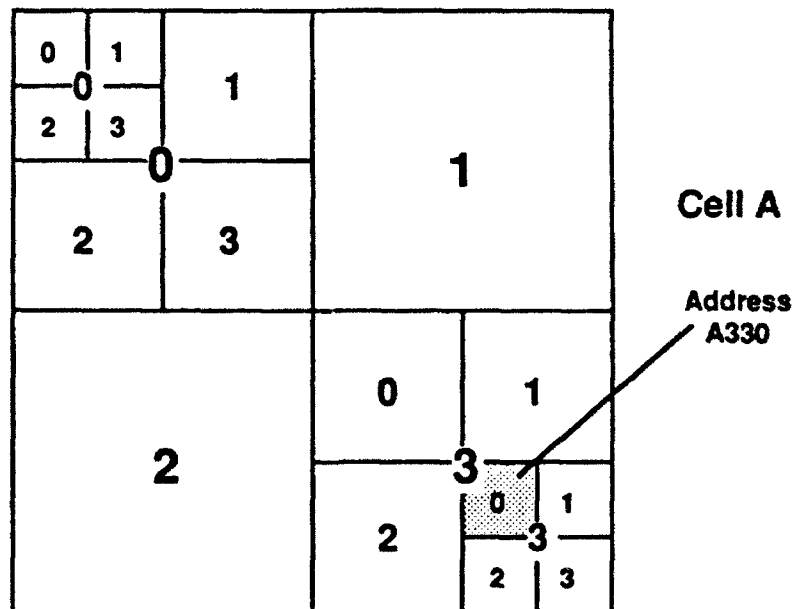
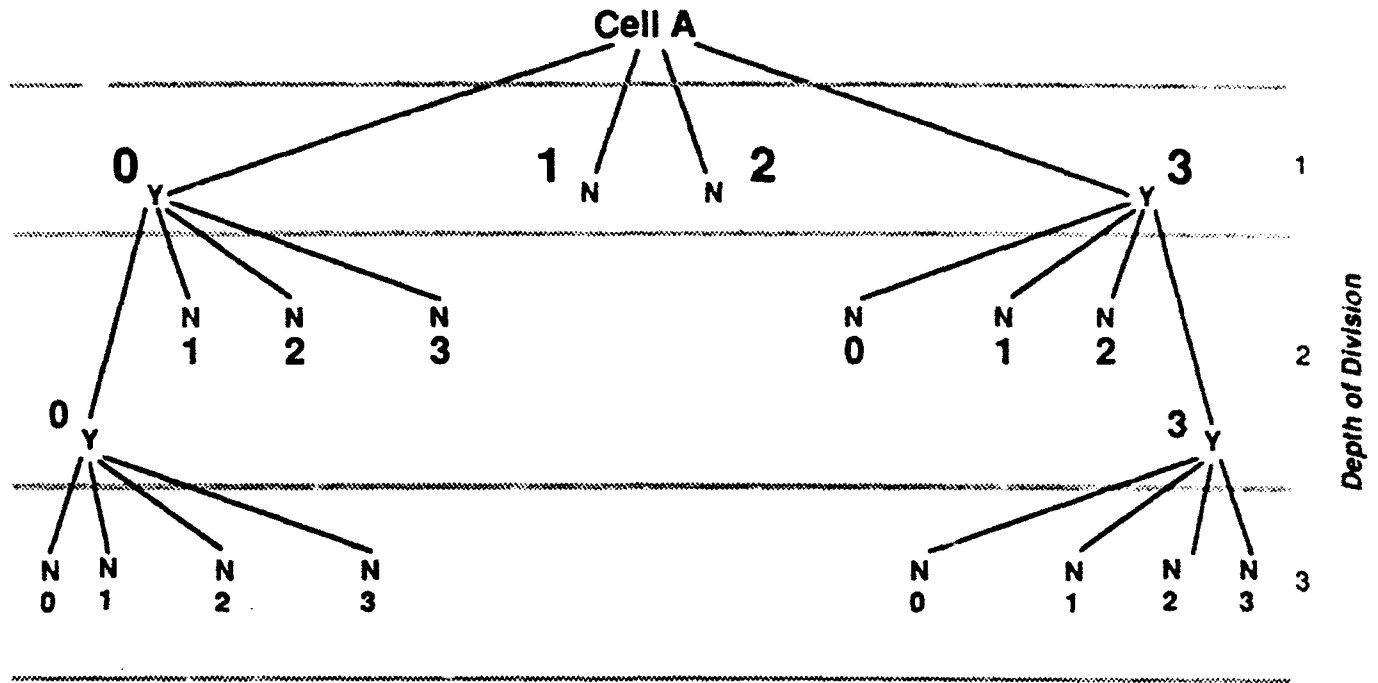


Figure 6. Quadtree Structure.

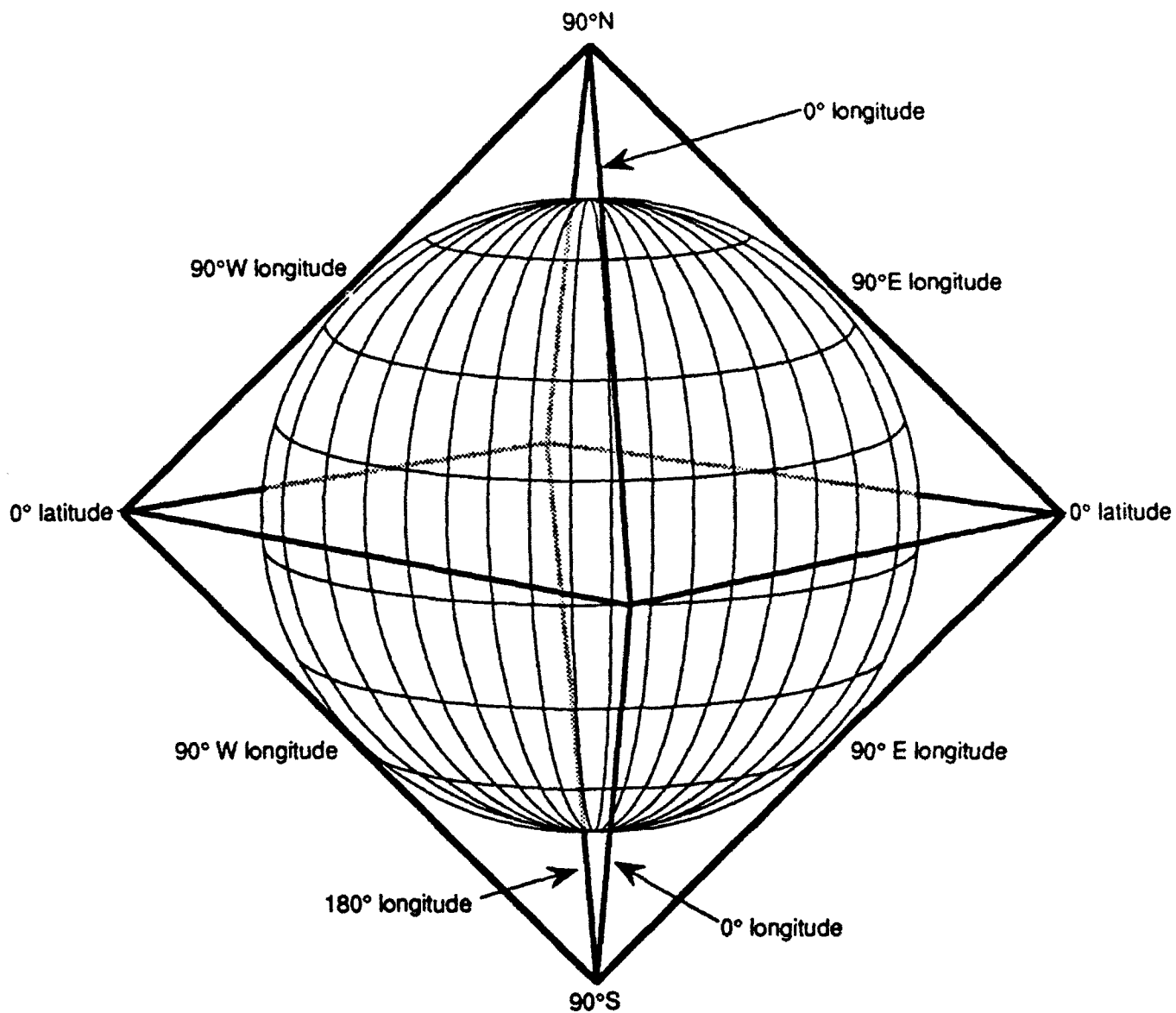
tiles are to apply to a solid, all the faces of the solid must be congruent and capable of recursive division into self-similar figures. Of the five polyhedra identified, the last four can be recursively divided.

Although the icosahedron has been extensively used in space-frame, or geodesic, construction, the octahedron (Figure 7) has been identified as the best candidate to serve as a model of the surface of the earth (Dutton, 1989). An octahedron can be recursively divided into self-similar triangles and also possesses other valuable properties. First, a line drawn between opposite long axis apexes of the octahedron will correspond to the axis of the earth. Once positioned in this manner, the solid is also aligned with the equator. Next, since each of the faces corresponds to one-quarter hemisphere, the solid can be rotated on its axis until its edges are in register along each 90° interval of longitude. Finally, each triangle base, including all those produced by recursion, will be along a line of latitude. Also, although at any given level the triangles produced are only approximately equal in area and shape, these violations are not nearly as severe as in any of the planar global systems (Goodchild and Siren, 1990). This property is very useful, since data density is a function of area. Maintaining tiles of approximately equal area anywhere on the globe will provide the most accurate method for obtaining units of equal data density. Recursive division of an octahedron yields a global, systematic, regular, adaptive tiling scheme.

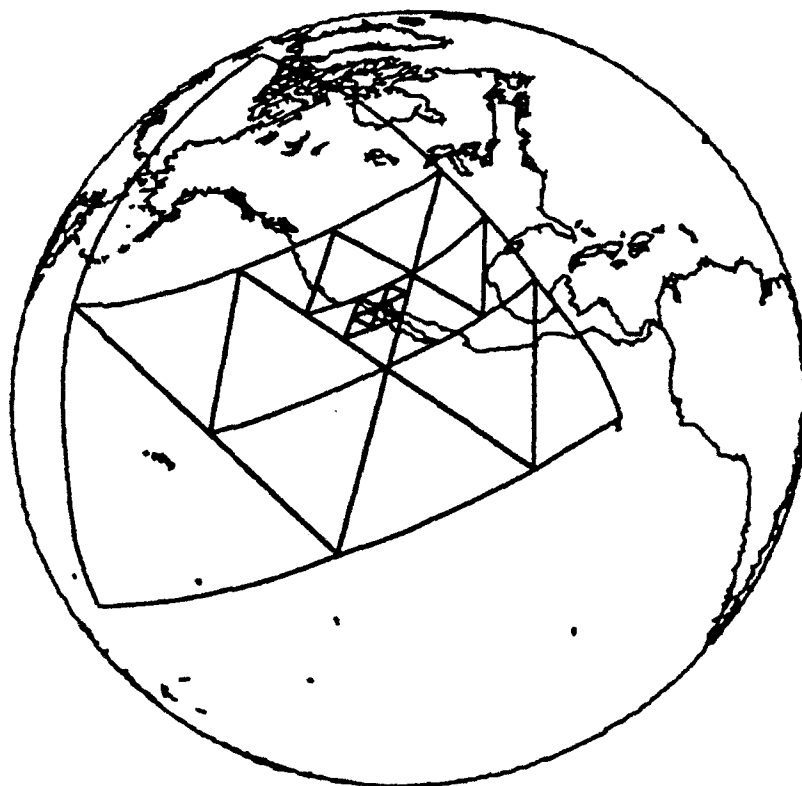
An examination of Table 2 reveals several properties of the octahedral model when divided recursively. At first glance it might seem that the number of triangles produced by recursive division, which quadruples at each level of division, would quickly produce an unmanageable amount of data. However, the area of the triangles that result from each division diminishes geometrically. After seven levels of division, the area of each tile is less than the average area of a 1° by 1° cell. An adaptive scheme, moreover, only allows (rather than requires) the earth to be tiled to a constant level over its entire surface. This property is relevant because the 1:1,000,000-scale ONC sheets as the source for the DCW cover land areas only. As data for the water areas becomes available, a recursive tiling scheme is the scheme that best accommodates the change in data density. Figure 8 illustrates a level 5 recursion focusing down to the area occupied by ONC sheet G-18 prototype area.

**Table 2. Length of Triangle Edges at Increasing Levels of Hemisphere**  
Hemisphere Radius Equals 6378 Km Along Equator  
and 6356 Km Along Meridian

Level	Length of triangle edge along Equator (longitude)	Length of triangle edge along Meridian (latitude)	Facet area (sq. km)
0	10018.5380 km	9983.8912 km	—
2	2504.6345 km	2495.9953 km	3,981,125
4	626.1586 km	623.9988 km	248,820
6	156.5397 km	155.9997 km	15,551
8	39.1349 km	38.9999 km	972
10	9.7837 km	9.7500 km	61



**Figure 7. Alignment of the Octahedron with the Earth. (Level 0 Triangles Are Equal to One-Quarter Hemisphere.)**



**Figure 8. Selective Tessellation of ONC G-18 Prototype Area: Level 0 Through 5 Triangles.**

## **Tiling Requirements and Goals**

For the DWC project, the characteristics of the tiling schemes described above must be evaluated in view of DCW requirements and goals. These requirements and goals are listed in relation to the various tiling schemes in Table 3 and discussed in the material below.

### **DCW Tiling Requirements**

First, the DCW Statement of Work (SOW; para. 1.1) states that the DCW database is to be global, and as a consequence the tiling scheme must be global.

Second, for the scheme to be effectively implemented, it must be capable of being integrated into the DCW production process. Of particular concern is the number of ONC map sheets that must be scanned and coded before tiling can proceed. The tiling scheme must permit partitioning to begin from a small number of edgematched sheets. The most desirable tiling scheme (from this standpoint) would allow each sheet to be integrated in the tiling scheme as the automation for it was completed.

Third, the selected tiling scheme must be in conformance with Vector Product Format (VPF) data structure. However, since VPF neither requires nor prohibits any given tiling scheme, conformance of the scheme to VPF is assured. In addition, VPF structure permits more than one tiling scheme to be used within a single database, since VPF structure accommodates multiple libraries.

Fourth, the tiling scheme must be applicable to and suitable for use with products at other scales. This requirement is stated explicitly in the SOW. The DCW database itself may be supplemented by different scale source materials.

Fifth, the tiling scheme must not make it necessary for a tile to occupy more than one CD (although a CD may contain more than one tile). In other words, only whole tiles can be written to each CD.

Sixth, the tiling scheme must support the results of the indexing studies, which are the subject of complementary research. In order to optimize data access and display, the tiling scheme must support or enhance the recommended indexing scheme.

### **DCW Tiling Goals**

The selected DCW tiling scheme should, insofar as possible, meet the seven goals listed below.

First, the tiling scheme should be compatible with existing tiling schemes. Since the tiling schemes currently supported by the Defense Mapping Agency are based on the graticule (although not all are global), the DCW tiling scheme should be compatible with graticule-based systems.

Second, the scheme should permit data to be returned to the user as rapidly as possible. Since the DCW is to be a general-purpose database with applications that access the CD directly, in order to optimize performance the tiling scheme will need to meet the requirements of the other components of the system, including those of the hardware configuration and display software.

Third, the scheme should prevent any single tile directory from containing more than the



Table 3. Requirements and Goals of Tiling Scheme

(a) Requirements

Tiling Scheme	Global in scope?	Can we integrate with production procedure?	Does it conform to VPF?	Usable with other scales?	Can we get whole tiles on CDs?	Does it support index study?
Non-systematic, fixed	Y	Y	Y	Y	Y	TBD (Indexing Study) ↑ ↓
Quasi-systematic, fixed	Y	Y	Y	Y	Y	
Systematic, fixed	Y	Y	Y	Y	Y	
Systematic, adaptive, irregular	Y	TBD (Proto 3 and 4) ↑ ↓	Y	Y	Y	
Systematic, adaptive, regular	Y		Y	Y	Y	
Nonplanar	Y		Y	Y	Y	

Note: Y=Yes.

(b) Goals

Tiling Scheme	Compatible with existing scheme?	Speed of access to tiles	Not exceed 40 files/dir	Does file size maximize throughput?	Is file size consistent?	Can tiles be adaptive?	Are tiles similar in size and shape?
Non-systematic, fixed	N	S	TBD (Indexing Study and Prototype 3) ↑ ↓	TBD (Prototype 3) ↑ ↓	N	N	N/N
Quasi-systematic, fixed	Y	S			N	Y	N/N
Systematic, fixed	Y	F			N	N	N/N
Systematic, adaptive, irregular	Y	M			Y	Y	N/N
Systematic, adaptive, regular	Y	F			Y	Y	N/N
Nonplanar	Y	F			Y	Y	Y/Y

Note: Y=Yes, N=No, S=Slow, M=Medium, F=Fast.

specified number of files. Forty files has been estimated as a practical CD-ROM limit. The indexing study and research currently being conducted with Prototype 3 may cause this number to be revised.

Fourth, consistent with results of the indexing studies, the tiling scheme should produce file sizes that maximize data throughput (throughput being the total time required to complete a query).

Fifth, to improve system predictability and optimize system performance, the tiles should be approximately the same size in terms of number of bytes. The level of predictability that can be provided by tiles of similar size will be assessed as part of the in-house evaluation of Prototype 3.

Sixth, to allow tiles to be approximately the same size in terms of bytes, the tiling scheme needs to be adaptive. Tiles of approximately the same data storage size are equivalent to tiles of approximately equal data density.

Seventh, to the extent possible, tiles should be of approximately the same shape and size for a given level of division. Achieving this goal will permit the most distortion-free and practical measurement of data density by means of area.

In addition to the requirements and goals listed above, the DCW tiling scheme may have to meet other criteria as a result of CD-ROM performance characteristics. These characteristics include, but are not limited to, changes in performance with changes in the number of directories, directory structure, and the sizes of the files located within each. These factors will directly affect the level of spatial division that will be acceptable to the CD and indirectly affect search and retrieval algorithms.

## **Current Efforts**

### **Prototype 2 Experience**

Valuable data density information was obtained from Prototype 2. Table 4 lists the data densities that were obtained from the Prototype 2 area of ONC sheet G-18. This prototype area included 28 1° by 1° cells, 20 of which contained data from 11 layers. Total data density varied from 600 bytes for the isogonic layer to over 1.2 megabytes for elevation. In any scheme based on data density, this wide range of values between the various layers will need to be taken into consideration.

### **Goals of Ongoing Effort**

**Prototype 3 Informational Goals.** Prototype 3 will provide a testbed for evaluating some DCW tiling issues. During the preparation of the Prototype 3 data for the CD-ROM pre-mastering tape, some tiling implementation issues have already been encountered. Tiling procedures developed for this effort include ARC Macro Language (AML) development and quality control procedures. A number of adjustments in these procedures will need to be made for use with Prototype 4.

Additional information on tiling issues will be available after the ESRI in-house testing and evaluation of the Prototype 3 database on the CD-ROM medium. Additional information about CD-ROM characteristics and the effects of tile size on the performance of the DCW will be available. A number of different tile sizes were placed on the CD-ROM to test the

Table 4. Approximate Data Densities for ONC G-18 in Prototype 2

Layer name	Feature type*	Total bytes by layer	Percent of total bytes	Bytes per cell for cells with data
Airtran	P	13,731	0	686
Country	L	1,183	0	59
Culture	PL	6,739	0	336
Drainage	PLA	378,645	17	18,932
Elevation	PLA	1,219,136	55	60,956
Isogonic	L	600	0	30
Landcov	LA	33,856	1	1,692
Poplace	PLA	135,270	6	6,763
Railroad	L	55,441	2	2,772
Road	L	233,006	10	11,650
Utility	L	135,181	6	6,759
		2,212,788	97	110,635

\*Feature type in the layer: P = point, L = line, A = area.

Total number of cells in ONC G-18 coverage: 28

Total number of 1° by 1° cells containing data (i.e., land surface): 20

Total area (sq. mi.) of ONC G-18 in Prototype 2: 100,124

Total area (sq. mi.) of 20 cells containing data: 79,000

effects of tile size on CD-ROM and system performance. Of particular interest is the effect of tile size on the time needed to draw a display on to the screen. It is hoped that these tests will result in a range of performance characteristics that reveals how tiles with different amounts of data affect the performance of other components of the system.

Prototype 3 will also permit the evaluation of the different types of tile addressing associated with each of the various tiling schemes being evaluated with Prototype 3 (nonplanar, fixed/regular, and systematic/irregular).

A number of other issues will be tested and evaluated which relate primarily to the indexing studies but have implications for tiling. These issues include a performance comparison between computing a tile address and searching an index of tile names. The effects of using various hierarchical directory structures for CD-ROM file seeks would be evaluated; the results might also affect tiling decisions.

The main function of Prototype 3 with regards to the tiling studies will be to narrow the number of options that need to be considered. Preliminary performance information on tile size and addressing will help the decision-making process move forward and identify those candidate schemes that will need to be evaluated more fully in the Prototype 4 effort

**Candidate Tiling Schemes for the DCW.** Given the program requirements and goals, three options were identified for further evaluation beginning with Prototype 3.

(1) The first, the quasi-systematic tiling scheme, is similar to the fixed tiling method used by the ONC sheets. It employs regular divisions along latitude and irregular divisions along longitude. The divisions along latitude might be the same as the  $8^\circ$  divisions within the ONC map series. The divisions along longitude would be determined by incrementally reducing the size of the partition until a set value of data density was reached. The most appropriate target value of data density will be identified after further exploration of CD-ROM characteristics and software functionality.

The advantages of this method are that it is conceptually straightforward and may in fact be a subset of the ONC distribution format. It has the added advantage of not presenting any extraordinary sheet edgematch problems. The disadvantages are that it is a fixed scheme, suffers from addressing shortcomings and, being constructed on a plane referenced in decimal degrees, poorly represents high-latitude areas. Since it is not systematic it may require considerable trial and error to bring into operation, and, although it yields tiles of a given data density, it does not yield tiles that are systematically related to each other.

(2) The second option is the systematic and adaptive subdivisions of a rectangle. Potentially, this system could begin with  $30^\circ$  by  $30^\circ$  squares at the highest level. Regular quad divisions of a  $30^\circ$  by  $30^\circ$  square would result in squares that are  $15^\circ$ ,  $7.5^\circ$ ,  $3.75^\circ$ ,  $1.875^\circ$ , and so on in size. Regular ninth divisions of a  $30^\circ$  square would result in tiles  $10^\circ$ ,  $1.111^\circ$ ,  $0.1234^\circ$  and so on in size. Regular quad divisions would be preferred, since the change in area with each division is smaller. An irregular scheme could also be tested that would divide the  $30^\circ$  square into quads for levels 1 and 2 then change in division to reflect a finer tuning toward a targeted data density or directory structure constraint.

The advantages of this second method lie in the systematic division of space, which would permit use of established tree search algorithms (Sedgewick, 1988; Samet, 1984). It is a global system, and the resulting tile boundaries could be in register to other referencing schemes, such as GEOREF. Also, implicit spatial relationships between tiles are derived for a level of effort similar to option 1 above. A prominent limitation is that since the scheme is planar, tile sizes and shapes will vary with position on the globe.

(3) The third option is the regular division of an octahedron. The advantage of this scheme over options 1 and 2 is that its construction on a solid approximately preserves area and shape for tiles at a given level of nesting anywhere on the globe. The disadvantages are that since it has never been implemented in a production environment, and the lead time necessary for testing, evaluation, and the development of implementation procedures could be considerable. This scheme also has the disadvantage of uncertain user acceptance, particularly if data from other sources is to be incorporated into the DCW tiling scheme. This is offset by the fact that VPF supports multiple libraries, each of which can contain a different tiling scheme. Efficient algorithms are already available that calculate a given triangle area, triangle address, and triangle vertices for any level of triangle based on the input of any longitude-latitude coordinates.

## Summary

The variability in the level of data density for the DCW requires that a tiling scheme be adaptive. In addition, in order to gain implicit spatial relationships and allow for rapid addressing, the scheme should be regular and systematic. For these reasons, regular divisions of a plane and regular divisions of an octahedron (options 2 and 3 above) will continue to be pursued as possible DCW tiling schemes.

## LITERATURE CITED

\_\_\_\_\_, 1988. *Product Specifications for Digital Feature Analysis Data (DFAD) Level 1-C and Level 3-C*, PS/1CK/200, PS/1CI/200 (Second ed.), St. Louis, MO: Defense Mapping Agency Aerospace Center.

\_\_\_\_\_, 1988. *Standard Index Chart*. ONC Index SIC-3 (Ed. 4), St. Louis, MO: Defense Mapping Agency Aerospace Center.

\_\_\_\_\_, 1986. *Datums, Ellipsoids, Grids, and Grid Reference Systems*, DMA TM 8358.1 (Prelim. ed.), Defense Mapping Agency Hydrographic/Topographic Center.

\_\_\_\_\_, 1986. *Product Specifications for Digital Terrain Elevation Data (DTED)*, PS/ICD/200, PS/ICF/200 (Second ed.), St. Louis, MO: Defense Mapping Agency Aerospace Center.

\_\_\_\_\_, 1986. *Which Map Is Best? Projections for World Maps*, Falls Church, VA: American Congress on Surveying and Mapping.

Broome, Frederick R., and Meixler, David B., 1990. The TIGER Data Base Structure, *Cartography and Geographic Information Systems* 17, No. 1, pp. 39-47.

Dutton, Geoffrey, 1989. Planetary modelling via Hierarchical Tessellation, *Auto-Carto 9, Proceedings Ninth International Symposium on Computer-Assisted Cartography*, ASPRS/ACSM, Baltimore, Md., April 2-7, pp. 87-96.

Goodchild, Michael, 1989. Optimal Tiling for Large Databases, *Auto-Carto 9, Proceedings Ninth International Symposium on Computer-Assisted Cartography*, ASPRS/ACSM, Baltimore, Md., April 2-7, pp. 444-451.

Goodchild, Michael, and Siren, Yang, 1990. A Hierarchical Spatial Data Structure for Global Geographic Information Systems (draft).

Peuquet, Donna, 1988. Issues Involved in Selecting Appropriate Data Models for Global Databases, *Building Databases for Global Science* (H. Mounsey and R. Tomlinson, Eds.), London: Taylor & Francis, pp. 66-78.

Robinson, Arthur H., and Sale, Randall D., 1969. *Elements of Cartography* (Third ed.), New York: John Wiley & Sons.

Samet, H., 1984. The Quadtree and Related Hierarchical Data Structures, *ACM Computing Surveys* 16, No. 2 (June), pp. 187-260.

Sedgewick, Robert, 1988. *Algorithms* (Second ed.), Reading, Mass: Addison-Wesley.